

- conventionally viewed as wrong, such as lying and stealing.
2. How does Mill reply to the objection that Epicureanism is a doctrine only worthy of swine?
 3. How does Mill distinguish between higher and lower pleasures?
 4. According to Mill, whose happiness must be considered?
 5. Carefully reconstruct Mill's proof of the Principle of Utility.

DISCUSSION QUESTIONS

1. Is happiness nothing more than pleasure, and the absence of pain? What do you think?



James Rachels

The Debate over Utilitarianism

James Rachels is university professor of philosophy at the University of Alabama of Birmingham. He is the author of *The End of Life: Euthanasia and Morality* (1986) and articles on the right to privacy, reverse discrimination, and the treatment of non-human animals. The reading is taken from his textbook *The Elements of Moral Philosophy* (1986).

Rachels presents the main objections to Utilitarianism and the replies given by defenders of Utilitarianism. His own view is that Utilitarianism is correct in telling us to consider the consequences of actions and in advising us to be impartial, but incorrect in ignoring other important moral considerations such as merit.

The utilitarian doctrine is that happiness is desirable, and the only thing desirable, as an end; all other things being desirable as means to that end.

John Stuart Mill, *Utilitarianism* (1861)

From James Rachels, *The Elements of Moral Philosophy*, (New York: McGraw-Hill Publishing Company, 1986).
Reprinted by permission of The McGraw-Hill Companies.

2. Does Mill convince you that the so-called higher pleasures are better than the lower ones? What about the person of experience who prefers the lower pleasures over the higher ones?
3. Mill says, "In the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility." Is this true or not?
4. Many commentators have thought that Mill's proof of the Principle of Utility is defective. Do you agree? If so, then what mistake or mistakes does he make? Is there any way to reformulate the proof so that it is not defective?

Man does not strive after happiness; only the Englishman does that.

Friedrich Nietzsche, *Twilight of the Idols* (1889)

THE RESILIENCE OF THE THEORY

Classical Utilitarianism—the theory defended by Bentham and Mill—can be summarized in three propositions:

First, actions are to be judged right or wrong solely in virtue of their consequences. Nothing else matters. Right actions are, simply, those that have the best consequences.

Second, in assessing consequences, the only thing that matters is the amount of happiness or unhappiness that is caused. Everything else is irrelevant. Thus right actions are those that produce the greatest balance of happiness over unhappiness.

Third, in calculating the happiness or unhappiness that will be caused, no one's happiness is to be counted as more important than anyone else's. Each person's welfare is equally important. As Mill put it in his *Utilitarianism*,

the happiness which forms the utilitarian standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator.



Digitized by the Internet Archive
in 2026 with funding from
CLIR DHC Grant 2026-2028

<https://archive.org/details/2174.15.06>

Thus right actions are those that produce the greatest possible balance of happiness over unhappiness, with each person's happiness counted as equally important.

The appeal of this theory to philosophers, economists, and others who theorize about human decision making has been enormous. The theory continues to be widely accepted, even though it has been challenged by a number of apparently devastating arguments. These anti-utilitarian arguments are so numerous, and so persuasive, that many have concluded the theory must be abandoned. But the remarkable thing is that so many have *not* abandoned it. Despite the arguments, a great many thinkers refuse to let the theory go. According to these contemporary utilitarians, the antiutilitarian arguments show only that the classical theory needs to be *modified*; they say the basic idea is correct and should be preserved, but recast into a more satisfactory form.

In what follows, we will examine some of these arguments against Utilitarianism, and consider whether the classical version of the theory may be revised satisfactorily to meet them. These arguments are of interest not only for the assessment of Utilitarianism but for their own sakes, as they raise some additional fundamental issues of moral philosophy.

IS HAPPINESS THE ONLY THING THAT MATTERS?

The question *What things are good?* is different from the question *What actions are right?* and Utilitarianism answers the second question by referring back to the first one. Right actions, it says, are the ones that produce the most good. But what is good? The classical utilitarian reply is: one thing, and one thing only, namely happiness. As Mill put it, "The utilitarian doctrine is that happiness is desirable, and the only thing desirable, as an end; all other things being desirable as means to that end."

The idea that happiness is the one ultimate good (and unhappiness the one ultimate evil) is known as Hedonism. Hedonism is a perennially popular theory that goes back at least as far as the ancient Greeks. It has always been an at-

tractive theory because of its beautiful simplicity, and because it expresses the intuitively plausible notion that things are good or bad only on account of the way they make us *feel*. Yet a little reflection reveals serious flaws in the theory. The flaws stand out when we consider examples like these:

1. A promising young pianist's hands are injured in an automobile accident so that she can no longer play. Why is this a bad thing for her? Hedonism would say it is bad because it causes her unhappiness. She will feel frustrated and upset whenever she thinks of what might have been, and *that* is her misfortune. But this way of explaining the misfortune seems to get things the wrong way around. It is not as though, by feeling unhappy, she has made an otherwise neutral situation into a bad one. On the contrary, her unhappiness is a rational response to a situation that is unfortunate. She could have had a career as a concert pianist, and now she cannot. *That* is the tragedy. We could not eliminate the tragedy just by getting her to cheer up.
2. You think someone is your friend, but really he ridicules you behind your back. No one ever tells you, so you never know. Is this situation unfortunate for you? Hedonism would have to say no, because you are never caused any unhappiness by the situation. Yet we do feel that there is something bad going on here. You *think* he is your friend, and you are "being made a fool," even though you are not aware of it and so suffer no unhappiness.

Both these examples make the same basic point. We value all sorts of things, including artistic creativity and friendship, for their own sakes. It makes us happy to have them, but only because we *already* think them good. (We do not think them good *because* they make us happy—this is what I meant when I said that Hedonism "gets things the wrong way around.") Therefore we think it a misfortune to lose them, independently of whether or not the loss is accompanied by unhappiness.

In this way, Hedonism misunderstands the nature of happiness. Happiness is not something

that is r
own sake
means of
is a resp
things t
dently a
friendshi
makes us
setting o
having fi
seeking f

Today,
of this.
donists.
have the
view wit
good and
losopher
to compi
as good
there are
ure, frie
that rich
world's s
have trie
things a
open que
are the o
goodness
Ideal Uti
the ques
we shoul
of peopl
ence Uti
this bool
these var
only in c
istic asst
largely b
have not
so by ur
essary pe

ARE CC
THAT M

The clai
however
most fur

that is recognized as good and sought for its own sake, with other things appreciated only as means of bringing it about. Instead, happiness is a response we have to the attainment of things that we recognize as goods, independently and in their own right. We think that friendship is a good thing, and so having friends makes us happy. That is very different from first setting out after happiness, then deciding that having friends might make us happy, and then seeking friends as a means to this end.

Today, most philosophers recognize the truth of this. There are not many contemporary hedonists. Those sympathetic to Utilitarianism have therefore sought a way to formulate their view without assuming a hedonistic account of good and evil. Some, such as the English philosopher G. E. Moore (1873–1958), have tried to compile short lists of things to be regarded as good in themselves. Moore suggested that there are three obvious intrinsic goods—pleasure, friendship, and aesthetic enjoyment—and that right actions are those that increase the world's supply of such things. Other utilitarians have tried to bypass the question of how many things are good in themselves, leaving it an open question and saying only that right actions are the ones that have the best results, *however* goodness is measured. (This is sometimes called Ideal Utilitarianism.) Still others try to bypass the question in another way, holding only that we should act so as to maximize the satisfaction of people's *preferences*. (This is called **Preference Utilitarianism**.) It is beyond the scope of this book to discuss the merits or demerits of these varieties of Utilitarianism. I mention them only in order to note that although the hedonistic assumption of the classical utilitarians has largely been rejected, contemporary utilitarians have not found it difficult to carry on. They do so by urging that Hedonism was never a necessary part of the theory in the first place.

ARE CONSEQUENCES ALL THAT MATTER?

The claim that only consequences matter is, however, a necessary part of Utilitarianism. The most fundamental idea underlying the theory is

that in order to determine whether an action would be right, we should look at *what will happen as a result of doing it*. If it were to turn out that some *other* matter is also important in determining rightness, then Utilitarianism would be undermined at its very foundation.

The most serious antiutilitarian arguments attack the theory at just this point: they urge that various other considerations, in addition to utility, are important in determining whether actions are right. We will look briefly at three such arguments.

1. *Justice*. Writing in the academic journal *Inquiry* in 1965, H. J. McCloskey asks us to consider the following case:

Suppose a utilitarian were visiting an area in which there was racial strife, and that, during his visit, a Negro rapes a white woman, and that race riots occur as a result of the crime, white mobs, with the connivance of the police, bashing and killing Negroes, etc. Suppose too that our utilitarian is in the area of the crime when it is committed such that his testimony would bring about the conviction of a particular Negro. If he knows that a quick arrest will stop the riots and lynchings, surely, as a utilitarian, he must conclude that he has a duty to bear false witness in order to bring about the punishment of an innocent person.

This is a fictitious example, but that makes no difference. The argument is only that *if* someone were in this position, then on utilitarian grounds he should bear false witness against the innocent person. This might have some bad consequences—the innocent man might be executed—but there would be enough good consequences to outweigh them: the riots and lynchings would be stopped. The best consequences would be achieved by lying; therefore, according to Utilitarianism, lying is the thing to do. But, the argument continues, it would be wrong to bring about the execution of the innocent man. Therefore, Utilitarianism, which implies it would be right, must be incorrect.

According to the critics of Utilitarianism, this argument illustrates one of the theory's most serious shortcomings: namely, that it is incompatible with the ideal of justice. Justice requires that we treat people fairly, according to their individual needs and merits. The innocent man

has done nothing wrong; he did not commit the rape and so he does not deserve to be punished for it. Therefore, punishing him would be unjust. The example illustrates how the demands of justice and the demands of utility can come into conflict, and so a theory that says utility is the whole story cannot be right.

2. *Rights*. Here is a case that is *not* fictitious; it is from the records of the U.S. Court of Appeals, Ninth Circuit (Southern District of California), 1963, in the case of *York v. Story*:

In October, 1958, appellant [Ms. Angelynn York] went to the police department of Chino for the purpose of filing charges in connection with an assault upon her. Appellee Ron Story, an officer of that police department, then acting under color of his authority as such, advised appellant that it was necessary to take photographs of her. Story then took appellant to a room in the police station, locked the door, and directed her to undress, which she did. Story then directed appellant to assume various indecent positions, and photographed her in those positions. These photographs were not made for any lawful purpose.

Appellant objected to undressing. She stated to Story that there was no need to take photographs of her in the nude, or in the positions she was directed to take, because the bruises would not show in any photograph. . . .

Later that month, Story advised appellant that the pictures did not come out and that he had destroyed them. Instead, Story circulated these photographs among the personnel of the Chino police department. In April, 1960, two other officers of that police department, appellee Louis Moreno and defendant Henry Grote, acting under color of their authority as such, and using police photographic equipment located at the police station made additional prints of the photographs taken by Story. Moreno and Grote then circulated these prints among the personnel of the Chino police department. . . .

Ms. York brought suit against these officers and won. Her legal rights had clearly been violated. But what of the *moral*ity of the officers' behavior?

Utilitarianism says that actions are defensible if they produce a favorable balance of happiness over unhappiness. This suggests that we consider the amount of unhappiness caused to Ms. York and compare it with the amount of pleasure taken in the photographs by Officer Story and

his cohorts. It is at least possible that more happiness than unhappiness was caused. In that case, the utilitarian conclusion apparently would be that their actions were morally all right. But this seems to be a perverse way to approach the case. Why should the pleasure afforded Story and his cohorts matter at all? Why should it even count? They had no right to treat Ms. York in that way, and the fact that they enjoyed doing so hardly seems a relevant defense.

To make the point even clearer, consider an (imaginary) related case. Suppose a Peeping Tom spied on Ms. York by peering through her bedroom window, and secretly took pictures of her undressed. Further suppose that he did this without ever being detected and that he used the photographs entirely for his own amusement, without showing them to anyone. Now under these circumstances, it seems clear that the *only* consequence of his action is an increase in his own happiness. No one else, including Ms. York, is caused any unhappiness at all. How, then, could Utilitarianism deny that the Peeping Tom's actions are right? But it is evident to moral common sense that they are not right. Thus, Utilitarianism appears to be an incorrect moral view.

The moral to be drawn from this argument is that Utilitarianism is at odds with the idea that people have *rights* that may not be trampled on merely because one anticipates good results. This is an extremely important notion, which explains why a great many philosophers have rejected Utilitarianism. In the above cases, it is Ms. York's right to privacy that is violated; but it would not be difficult to think of similar cases in which other rights are at issue—the right to freedom of religion, to free speech, or even the right to life itself. It may happen that good purposes are served, from time to time, by ignoring these rights. But we do not think that our rights *should* be set aside so easily. The notion of a personal right is not a utilitarian notion. Quite the reverse: it is a notion that places limits on how an individual may be treated, regardless of the good purposes that might be accomplished.

3. *Backward-Looking Reasons*. Suppose you have promised someone you will do something—say, you promised to meet him downtown this afternoon. But when the time comes

to go, you don't
some work and
should you do?
ity of getting y
outweighs the i
be caused. App
you might then
home. However
fact that you pr
you that you ca
if the consequen
were *great*—if,
just been strick
had to rush her
justified in brea
cannot overcor
fact that you
which says th
things that mat

There is an
learned from tl
anism vulnerab
because the on
the theory hol
rightness of acti
do with the fut
cern with cons
confine our att
result of our
think that can
have some imp
ised your friend
past, not the fut
to be made abo
to be an inadeq
cludes what w
considerations.

Once we unc
ples of backwa
easily to mind.
commit a crime
not be punishe
did you a favor
should now do
did something t
why you shoul
are all facts ab
determining ou
makes the past
cient for just th

to go, you don't want to do it—you need to do some work and would rather stay home. What should you do? Suppose you judge that the utility of getting your work accomplished slightly outweighs the inconvenience your friend would be caused. Appealing to the utilitarian standard, you might then conclude that it is right to stay home. However, this does not seem correct. The fact that you *promised* imposes an obligation on you that you cannot escape so easily. Of course, if the consequences of not breaking the promise were *great*—if, for example, your mother had just been stricken with a heart attack and you had to rush her to the hospital—you would be justified in breaking it. But a *small* gain in utility cannot overcome the obligation imposed by the fact that you promised. Thus Utilitarianism, which says that consequences are the only things that matter, seems mistaken.

There is an important general lesson to be learned from this argument. Why is Utilitarianism vulnerable to this sort of criticism? It is because the only kinds of considerations that the theory holds relevant to determining the rightness of actions are considerations having to do with the *future*. Because of its exclusive concern with consequences, Utilitarianism has us confine our attention to what *will happen* as a result of our actions. However, we normally think that considerations about the *past* also have some importance. The fact that you promised your friend to meet him is a fact about the past, not the future. Therefore, the general point to be made about Utilitarianism is that it seems to be an inadequate moral theory because it excludes what we might call backward-looking considerations.

Once we understand this point, other examples of backward-looking considerations come easily to mind. The fact that someone did not commit a crime is a good reason why he should not be punished. The fact that someone once did you a favor may be a good reason why you should now do him a favor. The fact that you did something to hurt someone may be a reason why you should now make it up to her. These are all facts about the past that are relevant to determining our obligations. But Utilitarianism makes the past irrelevant, and so it seems deficient for just that reason.

THE DEFENSE OF UTILITARIANISM

Taken together, the above arguments form an impressive indictment of Utilitarianism. The theory, which at first seemed so progressive and commonsensical, now seems indefensible: it is at odds with such fundamental moral notions as justice and individual rights, and seems unable to account for the place of backward-looking reasons in justifying conduct. The combined weight of these arguments has prompted many philosophers to abandon the theory altogether.

Many thinkers, however, continue to believe that Utilitarianism, in some form, is true. In reply to the arguments, three general defenses have been offered.

The First Line of Defense

The first line of defense is to point out that the examples used in the antiutilitarian arguments are unrealistic and do not describe situations that come up in the real world. Since Utilitarianism is designed as a guide for decision making in the situations we actually face, the fanciful examples are dismissed as irrelevant. . . .

The Second Line of Defense

The first line of defense contains more bluster than substance. While it can plausibly be maintained that *most* acts of false witness and the like have bad consequences in the real world, it cannot reasonably be asserted that *all* such acts have bad consequences. Surely, in at least some real-life cases, one can bring about good results by doing things that moral common sense condemns. Therefore, in at least some real-life cases Utilitarianism will come into conflict with common sense. Moreover, even if the antiutilitarian arguments had to rely exclusively on fictitious examples, those arguments would nevertheless retain their power; for showing that Utilitarianism has unacceptable consequences in hypothetical cases is a perfectly valid way of pointing up its theoretical defects. The first line of defense, then, is weak.

The second line of defense admits all this and proposes to 'save Utilitarianism by giving it a new formulation. In revising a theory to meet criticism, the trick is to identify precisely the

feature of the theory that is causing the trouble and to change *that*, leaving the rest of the theory undisturbed as much as possible.

The troublesome aspect of the theory was this: the classical version of Utilitarianism implied that *each individual action* is to be evaluated by reference to its own particular consequences. If on a certain occasion you are tempted to lie, whether it would be wrong is determined by the consequences of *that particular lie*. This, the theory's defenders said, is the point that causes all the trouble; even though we know that *in general* lying has bad consequences, it is obvious that sometimes particular acts of lying can have good consequences.

Therefore, the new version of Utilitarianism modifies the theory so that individual actions will no longer be judged by the Principle of Utility. Instead, *rules* will be established by reference to the principle, and individual acts will then be judged right or wrong by reference to the rules. This new version of the theory is called *Rule-Utilitarianism*, to contrast it with the original theory, now commonly called *Act-Utilitarianism*.

Rule-Utilitarianism has no difficulty coping with the three antiutilitarian arguments. An act-utilitarian, faced with the situation described by McCloskey, would be tempted to bear false witness against the innocent man because the consequences of *that particular act* would be good. But the rule-utilitarian would not reason in that way. He would first ask, "What *general rules of conduct* tend to promote the greatest happiness?" Suppose we imagine two societies, one in which the rule "Don't bear false witness against the innocent" is faithfully adhered to, and one in which this rule is not followed. In which society are people likely to be better off? Clearly, from the point of view of utility, the first society is preferable. Therefore, the rule against incriminating the innocent should be accepted, and *by appealing to this rule*, the rule-utilitarian concludes that the person in McCloskey's example should not testify against the innocent man.

Analogous arguments can be used to establish rules against violating people's rights, breaking promises, lying, and so on. We should accept such rules because following them, as a regular practice, promotes the general welfare. But once

having appealed to the Principle of Utility to establish the rules, we do not have to invoke the principle again to determine the rightness of particular actions. Individual actions are justified simply by appeal to the already-established rules.

Thus Rule-Utilitarianism cannot be convicted of violating our moral common sense, or of conflicting with ordinary ideas of justice, personal rights, and the rest. In shifting emphasis from the justification of acts to the justification of rules, the theory has been brought into line with our intuitive judgments to a remarkable degree.

The Third Line of Defense

Finally, a small group of contemporary utilitarians has had a very different response to the anti-utilitarian arguments. Those arguments point out that the classical theory is at odds with ordinary notions of justice, individual rights, and so on; to this, their response is, essentially, "So what?" In 1961 the Australian philosopher J. J. C. Smart published a monograph entitled *An Outline of a System of Utilitarian Ethics*; reflecting on his position in that book, Smart said:

Admittedly utilitarianism does have consequences which are incompatible with the common moral consciousness, but I tended to take the view "so much the worse for the common moral consciousness." That is, I was inclined to reject the common methodology of testing general ethical principles by seeing how they square with our feelings in particular instances.

Our moral common sense is, after all, not necessarily reliable. It may incorporate various irrational elements, including prejudices absorbed from our parents, our religion, and the general culture. Why should we simply assume that our feelings are always correct? And why should we reject a plausible, rational theory of ethics such as Utilitarianism simply because it conflicts with those feelings? Perhaps it is the feelings, not the theory, that should be discarded.

In light of this, consider again McCloskey's example of the person tempted to bear false witness. McCloskey argues that it would be wrong to have a man convicted of a crime he did not commit, because it would be unjust. But wait:

such a ju
enough,
who will
allowed t
might ho
like this,
tasteful.
curing th
(b) allow
ple, is it
option, b
On this
a perfect
need to
contrast,
sion of t
important
is, howev
which in
feelings'
what go
lenges us
tofore ta

WHAT IS INCORRECT

There is
can com
sequence
mote ha
admitted
Stuart M
are bene
standard
utilitaria
a part of
tion is
consider
knowlec
that are

If we
moral c
many ot
portant.
utilitaria
sense" c
that the
whites a

such a judgment serves *that man's* interests well enough, but what of the *other* innocent people who will be hurt if the rioting and lynchings are allowed to continue? What of them? Surely we might hope that we never have to face a situation like this, for the options are all extremely distasteful. But if we *must* choose between (a) securing the conviction of one innocent person and (b) allowing the deaths of several innocent people, is it so unreasonable to think that the first option, bad as it is, is preferable to the second?

On this way of thinking, Act-Utilitarianism is a perfectly defensible doctrine and does not need to be modified. Rule-Utilitarianism, by contrast, is an unnecessarily watered-down version of the theory, which gives rules a greater importance than they merit. Act-Utilitarianism is, however, recognized to be a radical doctrine which implies that many of our ordinary moral feelings may be mistaken. In this respect, it does what good philosophy always does—it challenges us to rethink matters that we have heretofore taken for granted.

WHAT IS CORRECT AND WHAT IS INCORRECT IN UTILITARIANISM

There is a sense in which no moral philosopher can completely reject Utilitarianism. The consequences of one's actions—whether they promote happiness, or cause misery—must be admitted by all to be extremely important. John Stuart Mill once remarked that, insofar as we are benevolent, we must accept the utilitarian standard; and he was surely right. Moreover, the utilitarian emphasis on impartiality must also be a part of any defensible moral theory. The question is whether these are the *only* kinds of considerations an adequate theory must acknowledge. Aren't there *other* considerations that are also important?

If we consult what Smart calls our "common moral consciousness," it seems that there are *many* other considerations that are morally important. . . . But I believe the radical act-utilitarians are right to warn us that "common sense" cannot be trusted. Many people once felt that there is an important difference between whites and blacks, so that the interests of whites

are somehow more important. Trusting the "common sense" of their day, they might have insisted that an adequate moral theory should accommodate this "fact." Today, no one worth listening to would say such a thing. But who knows how many *other* irrational prejudices are still a part of our moral common sense? At the end of his classic study of race relations, *An American Dilemma* (1944), the Swedish sociologist Gunnar Myrdal reminds us:

There must be still other countless errors of the same sort that no living man can yet detect, because of the fog within which our type of Western culture envelops us. Cultural influences have set up the assumptions about the mind, the body, and the universe with which we begin; pose the questions we ask; influence the facts we seek; determine the interpretation we give these facts; and direct our reaction to these interpretations and conclusions.

The strength of Utilitarianism is that it firmly resists "corruption" by possibly irrational elements. By sticking to the Principle of Utility as the *only* standard for judging right and wrong, it avoids all danger of incorporating into moral theory prejudices, feelings, and "intuitions" that have no rational basis.

The warning should be heeded. "Common sense" can, indeed, mislead us. At the same time, however, there might be at least some nonutilitarian considerations that an adequate theory *should* accept, because there is a rational basis for them. Consider, for example, the matter of what people deserve. A person who has worked hard in her job may deserve a promotion more than someone who has loafed, and it would be unjust for the loafer to be promoted first. This is a point that we would expect any fair-minded employer to acknowledge; we would all be indignant if we were passed over for promotion in favor of someone who had not worked as hard or as well as we. Now utilitarians might agree with this, and say that it can be explained by their theory—they might argue that it promotes the general welfare to encourage hard work by rewarding it. But this does not seem to be an adequate explanation of the importance of desert. The woman who worked harder has a superior claim to the promotion,

not because it promotes the general welfare for her to get it, but *because she has earned it*. The reason she should be promoted has to do with *her merits*. This does not appear to be the kind of consideration a utilitarian could admit.

Does this way of thinking express a mere prejudice, or does it have a rational basis? I believe it has a rational basis, although it is not one that utilitarians could accept. We ought to recognize individual desert as a reason for treating people in certain ways—for example, as a reason for promoting the woman who has worked harder—because that is the principal way we have of treating individuals as autonomous, responsible beings. If in fact people have the power to choose their own actions, in such a way that they are *responsible* for those actions and what results from them, then acknowledging their deserts is just a way of acknowledging their standing as autonomous individuals. In treating them as they deserve to be treated, we are responding to the way they have freely chosen to behave. Thus in some instances we will not treat everyone alike, because people are not just members of an undifferentiated crowd. Instead, they are individuals who, by their own choices, show themselves to deserve different kinds of responses. . . .

Immanuel Kant

The Categorical Imperative

Immanuel Kant (1724–1804), a German, was one of the most important philosophers of all time. He made significant contributions to all areas of philosophy. He wrote many books; the most important ones are *Critique of Pure Reason*, *Prolegomena*

From *The Moral Law: Kant's Groundwork of the Metaphysic of Morals*, trans. H. J. Paton (New York: Barnes & Noble, Inc., 1948).

REVIEW QUESTIONS

1. Rachels says that classical Utilitarianism can be summed up in three propositions. What are they?
2. Explain the problem with Hedonism. How do defenders of Utilitarianism respond to this problem?
3. What are the objections about justice, rights, and promises?
4. Distinguish between Rule- and Act-Utilitarianism. How does Rule-Utilitarianism reply to the objections?
5. What is the third line of defense?

DISCUSSION QUESTIONS

1. Smart's defense of Utilitarianism is to reject common moral beliefs when they conflict with Utilitarianism. Is this acceptable to you or not? Explain your answer.
2. A utilitarian is supposed to give moral consideration to all concerned. Who must be considered? What about nonhuman animals? How about lakes and streams?
3. Rachels claims that merit should be given moral consideration independent of utility. Do you agree?

to *All Future Metaphysics*, *Critique of Practical Reason*, *Critique of Judgment*, and *The Foundations of the Metaphysics of Morals*, from which the reading is taken.

Kant believes that our moral duty can be formulated in one supreme rule, the *Categorical Imperative*, from which all our duties can be derived. Although he says that there is just one rule, he gives different versions of it, and two of them seem to be distinct. He arrives at the supreme rule or rules by considering the nature of the good will and duty.

THE GOOD WILL

It is impossible to conceive anything at all in the world, or even out of it, which can be taken as

good will
telligence
the mind
olution,
temper
ble in n
tremely
which h
and whi
applied
same wi
even he
contentr
name of
consequ
good wi
the min
tion—m
ends; no
spectato
the unin
no toucl
sequent
dispense
be happ

Some
will itse
They ha
worth, l
sets a l
rightly
them as
tions an
ion are
even se
of a per
describe
uncond
the anc
good wi
the very
merely
more a
have tal

THE G
A good
fects or

good without qualification, except a *good will*. Intelligence, wit, judgment, and any other *talents* of the mind we may care to name, or courage, resolution, and constancy of purpose, as qualities of *temperament*, are without doubt good and desirable in many respects; but they can also be extremely bad and hurtful when the will is not good which has to make use of these gifts of nature, and which for this reason has the term '*character*' applied to its peculiar quality. It is exactly the same with *gifts of fortune*. Power, wealth, honour, even health and that complete well-being and contentment with one's state which goes by the name of '*happiness*', produce boldness, and as a consequence often over-boldness as well, unless a good will is present by which their influence on the mind—and so too the whole principle of action—may be corrected and adjusted to universal ends; not to mention that a rational and impartial spectator can never feel approval in contemplating the uninterrupted prosperity of a being graced by no touch of a pure and good will, and that consequently a good will seems to constitute the indispensable condition of our very worthiness to be happy.

Some qualities are even helpful to this good will itself and can make its task very much easier. They have none the less no inner unconditioned worth, but rather presuppose a good will which sets a limit to the esteem in which they are rightly held and does not permit us to regard them as absolutely good. Moderation in affections and passions, self-control, and sober reflexion are not only good in many respects: they may even seem to constitute part of the *inner* worth of a person. Yet they are far from being properly described as good without qualification (however unconditionally they have been commended by the ancients). For without the principles of a good will they may become exceedingly bad; and the very coolness of a scoundrel makes him, not merely more dangerous, but also immediately more abominable in our eyes than we should have taken him to be without it.

THE GOOD WILL AND ITS RESULTS

A good will is not good because of what it effects or accomplishes—because of its fitness for

attaining some proposed end: it is good through its willing alone—that is, good in itself. Considered in itself it is to be esteemed beyond comparison as far higher than anything it could ever bring about merely in order to favour some inclination or, if you like, the sum total of inclinations. Even if, by some special disfavour of destiny or by the niggardly endowment of step-motherly nature, this will is entirely lacking in power to carry out its intentions; if by its utmost effort it still accomplishes nothing, and only good will is left (not, admittedly, as a mere wish, but as the straining of every means so far as they are in our control); even then it would still shine like a jewel for its own sake as something which has its full value in itself. Its usefulness or fruitlessness can neither add to, nor subtract from, this value. Its usefulness would be merely, as it were, the setting which enables us to handle it better in our ordinary dealings or to attract the attention of those not yet sufficiently expert, but not to commend it to experts or to determine its value. . . .

THE GOOD WILL AND DUTY

We have now to elucidate the concept of a will estimable in itself and good apart from any further end. This concept, which is already present in a sound natural understanding and requires not so much to be taught as merely to be clarified, always holds the highest place in estimating the total worth of our actions and constitutes the condition of all the rest. We will therefore take up the concept of *duty*, which includes that of a good will, exposed, however, to certain subjective limitations and obstacles. These, so far from hiding a good will or disguising it, rather bring it out by contrast and make it shine forth more brightly.

THE MOTIVE OF DUTY

I will here pass over all actions already recognized as contrary to duty, however useful they may be with a view to this or that end; for about these the question does not even arise whether they could have been done *for the sake of duty* inasmuch as they are directly opposed to it. I

will also set aside actions which in fact accord with duty, yet for which men have *no immediate inclination*, but perform them because impelled to do so by some other inclination. For there it is easy to decide whether the action which accords with duty has been done *from duty* or from some purpose of self-interest. This distinction is far more difficult to perceive when the action accords with duty and the subject has in addition an *immediate* inclination to the action. For example, it certainly accords with duty that a grocer should not overcharge his inexperienced customer; and where there is much competition a sensible shopkeeper refrains from so doing and keeps to a fixed and general price for everybody so that a child can buy from him just as well as anyone else. Thus people are served *honestly*; but this is not nearly enough to justify us in believing that the shopkeeper has acted in this way from duty or from principles of fair dealing; his interests required him to do so. We cannot assume him to have in addition an immediate inclination towards his customers, leading him, as it were out of love, to give no man preference over another in the matter of price. Thus the action was done neither from duty nor from immediate inclination, but solely from purposes of self-interest.

On the other hand, to preserve one's life is a duty, and besides this every one has also an immediate inclination to do so. But on account of this the often anxious precautions taken by the greater part of mankind for this purpose have no inner worth, and the maxim of their action is without moral content. They do protect their lives *in conformity with duty*, but not *from the motive of duty*. When on the contrary, disappointments and hopeless misery have quite taken away the taste for life; when a wretched man, strong in soul and more angered at his fate than faint-hearted or cast down, longs for death and still preserves his life without loving it—not from inclination or fear but from duty; then indeed his maxim has a moral content.

To help others where one can is a duty, and besides this there are many spirits of so sympathetic a temper that, without any further motive of vanity or self-interest, they find an inner pleasure in spreading happiness around them and can take delight in the contentment of others as their

own work. Yet I maintain that in such a case an action of this kind, however right and however amiable it may be, has still no genuinely moral worth. It stands on the same footing as other inclinations—for example, the inclination for honour, which if fortunate enough to hit on something beneficial and right and consequently honourable, deserves praise and encouragement, but not esteem; for its maxim lacks moral content, namely, the performance of such actions, not from inclination, but *from duty*. Suppose then that the mind of this friend of man were overclouded by sorrows of his own which extinguished all sympathy with the fate of others, but that he still had power to help those in distress, though no longer stirred by the need of others because sufficiently occupied with his own; and suppose that, when no longer moved by any inclination, he tears himself out of this deadly insensibility and does the action without any inclination for the sake of duty alone; then for the first time his action has its genuine moral worth. Still further: if nature had implanted little sympathy in this or that man's heart; if (being in other respects an honest fellow) he were cold in temperament and indifferent to the sufferings of others—perhaps because, being endowed with the special gift of patience and robust endurance in his own sufferings, he assumed the like in others or even demanded it; if such a man (who would in truth not be the worth product of nature) were not exactly fashioned by her to be a philanthropist, would he not still find in himself a source from which he might draw a worth far higher than any that a good-natured temperament can have? Assuredly he would. It is precisely in this that the worth of character begins to show—a moral worth and beyond all comparison the highest—namely, that he does good, not from inclination, but from duty. . . .

Thus the moral worth of an action does not depend on the result expected from it, and so too does not depend on any principle of action that needs to borrow its motive from this expected result. For all these results (agreeable states and even the promotion of happiness in others) could have been brought about by other causes as well, and consequently their production did not require the will of a rational being, in which, however, the highest and uncondi-

tioned good nothing but *admittedly* far as it, a ground det that preemi good which ing on this merely from

THE CATE

But what ki which, even expected from is to be call ification? Si inducement quence of (is left but tl law as such as its princ act except i my maxim bare confor out having ticular acti principle, a be everywl merical coi kind also practical ju said princij

When I general, I c contain—u ceive a *cat* what it cor imperative maxim¹ sh law, as we limit it, the maxim has a law as su the impera

There is imperative through wh it should be

tioned good can alone be found. Therefore nothing but the *idea of the law in itself, which admittedly is present only in a rational being*—so far as it, and not an expected result, is the ground determining the will—can constitute that preeminent good which we call moral, a good which is already present in the person acting on this idea and has not to be awaited merely from the result.

THE CATEGORICAL IMPERATIVE

But what kind of law can this be the thought of which, even without regard to the results expected from it, has to determine the will if this is to be called good absolutely and without qualification? Since I have robbed the will of every inducement that might arise for it as a consequence of obeying any particular law, nothing is left but the conformity of actions to universal law as such, and this alone must serve the will as its principle. That is to say, I ought never to act except in such a way that *I can also will that my maxim should become a universal law*. Here bare conformity to universal law as such (without having as its base any law prescribing particular actions) is what serves the will as its principle, and must so serve it if duty is not to be everywhere an empty delusion and a chimerical concept. The ordinary reason of mankind also agrees with this completely in its practical judgements and always has the aforesaid principle before its eyes. . . .

When I conceive a *hypothetical imperative* in general, I do not know beforehand what it will contain—until its condition is given. But if I conceive a *categorical imperative*, I know at once what it contains. For since besides the law this imperative contains only the necessity that our maxim¹ should conform to this law, while the law, as we have seen, contains no condition to limit it, there remains nothing over to which the maxim has to conform except the universality of a law as such; and it is this conformity alone that the imperative properly asserts to be necessary.

There is therefore only a single categorical imperative and it is this: '*Act only on that maxim through which you can at the same time will that it should become a universal law*'.

Now if all imperatives of duty can be derived from this one imperative as their principle, then even although we leave it unsettled whether what we call duty may not be an empty concept, we shall still be able to show at least what we understand by it and what the concept means. . . .

ILLUSTRATIONS

We will now enumerate a few duties, following their customary division into duties towards self and duties towards others and into perfect and imperfect duties.²

1. A man feels sick of life as the result of a series of misfortunes that has mounted to the point of despair, but he is still so far in possession of his reason as to ask himself whether taking his own life may not be contrary to his duty to himself. He now applies the test 'Can the maxim of my action really become a universal law of nature?' His maxim is 'From self-love I make it my principle to shorten my life if its continuance threatens more evil than it promises pleasure'. The only further question to ask is whether this principle of self-love can become a universal law of nature. It is then seen at once that a system of nature by whose law the very same feeling whose function (*Bestimmung*) is to stimulate the furtherance of life should actually destroy life would contradict itself and consequently could not subsist as a system of nature. Hence this maxim cannot possibly hold as a universal law of nature and is therefore entirely opposed to the supreme principle of all duty.

2. Another finds himself driven to borrowing money because of need. He well knows that he will not be able to pay it back; but he sees too that he will get no loan unless he gives a firm promise to pay it back within a fixed time. He is inclined to make such a promise; but he has still enough conscience to ask 'Is it not unlawful and contrary to duty to get out of difficulties in this way?' Supposing, however, he did resolve to do so, the maxim of his action would run thus: 'Whenever I believe myself short of money, I will borrow money and promise to pay it back, though I know that this will never be done'. Now this principle of self-love or personal advantage is perhaps quite compatible

with my own entire future welfare; only there remains the question 'Is it right?' I therefore transform the demand of self-love into a universal law and frame my question thus: 'How would things stand if my maxim became a universal law?' I then see straight away that this maxim can never rank as a universal law of nature and be self-consistent, but must necessarily contradict itself. For the universality of a law that every one believing himself to be in need can make any promise he pleases with the intention not to keep it would make promising, and the very purpose of promising, itself impossible, since no one would believe he was being promised anything, but would laugh at utterances of this kind as empty shams.

3. A third finds in himself a talent whose cultivation would make him a useful man for all sorts of purposes. But he sees himself in comfortable circumstances, and he prefers to give himself up to pleasure rather than to bother about increasing and improving his fortunate natural aptitudes. Yet he asks himself further 'Does my maxim of neglecting my natural gifts, besides agreeing in itself with my tendency to indulgence, agree also with what is called duty?' He then sees that a system of nature could indeed always subsist under such a universal law, although (like the South Sea Islanders) every man should let his talents rust and should be bent on devoting his life solely to idleness, indulgence, procreation, and, in a word, to enjoyment. Only he cannot possibly will that this should become a universal law of nature or should be implanted in us as such a law by a natural instinct. For as a rational being he necessarily wills that all his powers should be developed, since they serve him, and are given him, for all sorts of possible ends.

4. Yet a fourth is himself flourishing, but he sees others who have to struggle with great hardships (and whom he could easily help); and he thinks 'What does it matter to me? Let every one be as happy as Heaven wills or as he can make himself; I won't deprive him of anything; I won't even envy him; only I have no wish to contribute anything to his well-being or to his support in distress!' Now admittedly if such an attitude were a universal law of nature, mankind could get on perfectly well—better no doubt

than if everybody prates about sympathy and goodwill, and even takes pains, on occasion, to practise them, but on the other hand cheats where he can, traffics in human rights, or violates them in other ways. But although it is possible that a universal law of nature could subsist in harmony with this maxim, yet it is impossible to will that such a principle should hold everywhere as a law of nature. For a will which decided in this way would be in conflict with itself, since many a situation might arise in which the man needed love and sympathy from others, and in which, by such a law of nature sprung from his own will, he would rob himself of all hope of the help he wants for himself. . . .

THE FORMULA OF THE END IN ITSELF

The will is conceived as a power of determining oneself to action *in accordance with the idea of certain laws*. And such a power can be found only in rational beings. Now what serves the will as a subjective ground of its self-determination is an *end*; and this, if it is given by reason alone, must be equally valid for all rational beings. What, on the other hand, contains merely the ground of the possibility of an action whose effect is an end is called a *means*. . . .

Now I say that man, and in general every rational being, exists as an end in himself, *not merely as a means* for arbitrary use by this or that will: he must in all his actions, whether they are directed to himself or to other rational beings, always be viewed *at the same time as an end*. All the objects of inclination have only a conditioned value; for if there were not these inclinations and the needs grounded on them, their object would be valueless. Inclinations themselves, as sources of needs, are so far from having an absolute value to make them desirable for their own sake that it must rather be the universal wish of every rational being to be wholly free from them. Thus the value of all objects that can be *produced* by our action is always conditioned. Beings whose existence depends, not on our will, but on nature, have none the less, if they are non-rational beings, only a relative value as means and are consequently

called things; in hand, are cal ready marks that is, as soi merely as a to that exter of them (and therefore, an existence as for us: they whose exist an end such other end to means; for absolute valu all value w gent—then found for re

If then the ciple and— cerned—a ce that from th sarily an enu in itself it fo and consequ The ground exists as an a man nece it is therefo man actions other ration the same ra me; hence it ciple, from v it must be p The practic lows: Act in manity, whu person of an always at th

Endnotes

1. A maxim is distinguished

called *things*. Rational beings, on the other hand, are called *persons* because their nature already marks them out as ends in themselves—that is, as something which ought not to be used merely as a means—and consequently imposes to that extent a limit on all arbitrary treatment of them (and is an object of reverence). Persons, therefore, are not merely subjective ends whose existence as an object of our actions has a value for us: they are *objective ends*—that is, things whose existence is in itself an end, and indeed an end such that in its place we can put no other end to which they should serve *simply* as means; for unless this is so, nothing at all of *absolute* value would be found anywhere. But if all value were conditioned—that is, contingent—then no supreme principle could be found for reason at all.

If then there is to be a supreme practical principle and—so far as the human will is concerned—a categorical imperative, it must be such that from the idea of something which is necessarily an end for every one because it is an *end in itself* it forms an *objective* principle of the will and consequently can serve as a practical law. The ground of this principle is: *Rational nature exists as an end in itself*. This is the way in which a man necessarily conceives his own existence: it is therefore so far a *subjective* principle of human actions. But it is also the way in which every other rational being conceives his existence on the same rational ground which is valid also for me; hence it is at the same time an *objective* principle, from which, as a supreme practical ground, it must be possible to derive all laws for the will. The practical imperative will therefore be as follows: *Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end.* . . .

Endnotes

1. A *maxim* is a subjective principle of action and must be distinguished from an *objective principle*—namely, a

practical law. The former contains a practical rule determined by reason in accordance with the conditions of the subject (often his ignorance or again his inclinations); it is thus a principle on which the subject *acts*. A law, on the other hand, is an objective principle valid for every rational being; and it is a principle on which he *ought to act*—that is, an imperative.

2. It should be noted that I reserve my division of duties entirely for a future *Metaphysic of Morals* and that my present division is therefore put forward as arbitrary (merely for the purpose of arranging my examples). Further, I understand here by a perfect duty one which allows no exception in the interests of inclination, and so I recognize among *perfect duties*, not only outer ones, but also inner. This is contrary to the accepted usage of the schools, but I do not intend to justify it here, since for my purpose it is all one whether this point is conceded or not.

REVIEW QUESTIONS

1. Explain Kant's account of the good will.
2. Distinguish between hypothetical and categorical imperatives.
3. State the first formulation of the Categorical Imperative (using the notion of a universal law) and explain how Kant uses this rule to derive some specific duties towards self and others.
4. State the second version of the Categorical Imperative (using the language of means and end), and explain it.

DISCUSSION QUESTIONS

1. Are the two different versions of the Categorical Imperative just different expressions of one basic rule, or are they two different rules? Defend your view.
2. Kant claims that an action that is not done from the motive of duty has no moral worth. Do you agree or not? If not, give some counter-examples.
3. Some commentators think that the Categorical Imperative (particularly the first formulation) can be used to justify nonmoral or immoral actions. Is this a good criticism?

